

ROBIN HANSON

¿Qué pasaría si transfiriéramos nuestros cerebros a las computadoras?

Algún día puede que tengamos robots tan inteligentes como personas, inteligencia artificial, IA. ¿Cómo podría suceder eso? Un camino es que vamos a seguir acumulando mejor software, como lo hemos estado haciendo desde hace 70 años. A ritmos de progreso anteriores, eso podría tomar siglos. Algunos dicen que sucederá mucho más rápido, a medida que descubramos grandes, nuevas y poderosas teorías de la inteligencia. Soy escéptico.

Pero sobre un tercer escenario, es de lo que hablaré hoy. La idea es portar el software desde el cerebro humano. Para hacer esto, necesitaremos tres tecnologías que sean lo suficientemente buenas y ninguna de ellas existe aún. Primero que todo, necesitaremos muchas computadoras baratas, rápidas y paralelas. Segundo, necesitaremos escanear cerebros de individuos humanos, en fino detalle espacial y químico, para ver exactamente en qué lugar están conectadas las células y a qué y qué tipo. Y tercero, necesitaremos modelos de computadora de cómo funciona cada tipo de neurona... tomar las señales de entrada, cambiar el estado de intervalo y enviar señales de salida. Si tenemos suficientes modelos buenos de todos los tipos de neuronas y un modelo del cerebro suficientemente bueno, podemos juntarlo para hacer un modelo, suficientemente bueno de todo un cerebro y ese modelo tendría el mismo comportamiento del original. Así que, si le hablan a eso, podría responder. Si le piden que haga cosas, podría hacerlas. Y si pudiéramos hacer eso, todo cambiaría.

La gente ha estado hablando de esta idea durante décadas, bajo el nombre de "*uploads*". Los llamaré "*ems*". Cuando hablan de ello, dicen: "¿Es esto siquiera posible? Si hicieras uno, ¿sería consciente? ¿O es solo una máquina vacía? Si hicieras uno de mí, ¿sería yo o alguien más?".

Todas estas son preguntas fascinantes, que voy a ignorar... porque veo una pregunta olvidada: ¿Qué pasaría realmente? Me obsesioné con esta pregunta. Pasé cuatro años tratando de analizarla, utilizando herramientas académicas estándar, para descifrar qué pasaría y estoy aquí para contarles lo que encontré. Pero tengan cuidado, no estoy ofreciendo inspiración, estoy ofreciendo análisis. Veo mi trabajo como el contarles, lo que es más probable que suceda, si hicimos muy poco para evitarlo. Si no están al menos un poco perturbados por lo que les diga aquí, simplemente no están prestando la atención.

Bueno, lo primero que puedo decir es que los *ems* pasan la mayor parte de su vida en la realidad virtual. A esto podría parecerse si utilizaran la realidad virtual. Y esto es lo que podrían ver: la luz del sol que destella en el agua, podrían oír las gaviotas volando, podrían incluso sentir el viento en sus mejillas u oler el agua de mar con un hardware avanzado. Ahora, si pasaran mucho tiempo aquí, podrían querer un panel de mandos,

donde pudieran hacer cosas como una llamada telefónica, pasar a un nuevo mundo virtual, revisar su cuenta bancaria.

Mientras que es así como les parecería a Uds. la realidad virtual, así es como un **em** se vería en realidad virtual. Es hardware de computadora, ubicado en un servidor en alguna parte. Aún así, podría ver y experimentar lo mismo. Pero algunas cosas son diferentes para los **ems**. Primero, mientras que Uds. probablemente siempre notarán que la realidad virtual no es del todo verdadera, un **em**, puede sentirla tan real como ahora es esta sala para Uds., o como cualquier cosa que uno siente. Los **ems** también tienen algunas posibilidades más de acción.

Por ejemplo, nuestra mente siempre corre a la misma velocidad, pero un **em** puede agregar o quitar hardware para operar más rápido o lento y, por lo tanto, si el mundo alrededor parece ir demasiado rápido, simplemente pueden acelerar sus mentes y el mundo que les rodea parecería ralentizarse.

Además, un **em** puede hacer una copia de sí mismo en ese momento. Esta copia recordaría todo igual y si comienza con la misma velocidad --observando a la misma velocidad-- podría incluso necesitar que le dijeran: "Tú eres la copia". Y un **em** podría archivar copias de archivos y con suficientes archivos, un **em** puede ser inmortal; en teoría, aunque no en la práctica por lo general. Y un **em** puede trasladar su cerebro, la computadora que representa su cerebro, de una ubicación física a otra. De hecho, los **ems** pueden viajar por el mundo a la velocidad de la luz y al moverse a una nueva ubicación, pueden interactuar más rápidamente con **ems** cerca de esa nueva ubicación.

Hasta ahora, he estado hablando de lo que los **ems** pueden hacer. ¿Qué deciden hacer los **ems**? Para entender eso, necesitaremos entender tres hechos claves. **Primero...** por definición los **ems** emulan lo que el humano haría en la misma situación. Así que, su vida y comportamiento es muy humano. Son principalmente diferentes porque están viviendo en un mundo diferente. **Segundo**, los **ems** necesitan recursos reales para sobrevivir. Uds. necesitan comida y refugio o de lo contrario morirían. Del mismo modo, los **ems** necesitan hardware, energía, refrigeración, pues de otra manera no pueden existir. Por cada minuto subjetivo que experimenta un **em**, alguien, en general ese **em**, tuvo que trabajar para pagar por ello. **Tercero**, los **ems** son pobres.

La población de **ems** puede crecer más rápido que la economía de los **ems**, de modo que los salarios se reducen a los niveles de subsistencia de **ems**. Eso significa que los **ems** deben trabajar la mayoría del tiempo. Así que esto significa, que esto es lo que los **ems** generalmente ven: hermosos y lujosos... pero al final escritorios. Trabajan la mayoría del tiempo. Podría pensarse que un escenario de salario de subsistencia es exótico y raro, pero en realidad es el caso habitual en la historia humana y es como casi todos los animales salvajes han vivido, así que sabemos lo que hacen los humanos en

esta situación. Básicamente, los humanos hacen lo que se necesita para sobrevivir y esto es lo que me permite decir tanto acerca del mundo de los *ems*. Cuando las criaturas son ricas, como Uds., hay que saber mucho acerca de lo que ellos quieren para averiguar lo que hacen. Cuando las criaturas son pobres, la mayoría hace lo necesario para sobrevivir.

Hemos hablado del mundo de los *ems*, desde la perspectiva de los *ems*. Ahora, demos un paso atrás y observemos todo su mundo. Primero, el mundo de los *ems* crece mucho más rápido que el nuestro, cerca de cien veces más rápido. Así que la misma cantidad de cambio que experimentaríamos en un siglo o dos, ellos lo experimentarían en un año o dos. Y no estoy dispuesto a proyectar esta época mucho más allá de eso, porque posiblemente para entonces, algo más sucederá, no sé qué. Segundo, la emulación típica se ejecuta aún más rápido, aproximadamente mil veces la velocidad humana. Así que, ellos experimentan miles de años en ese año o dos, y para ellos el mundo que les rodea cambia más lentamente, de lo que el mundo parece cambiar para Uds. Tercero, los *ems* se hacinan en un número pequeño de ciudades muy pobladas. Esto no es solo cómo se ven a sí mismas en la realidad virtual, también es cómo realmente están hacinadas físicamente. Así que, a velocidad de *ems*, los viajes físicos parecen dolorosamente lentos, por lo que la mayoría de las ciudades son autosuficientes, las guerra cibernéticas, y la mayor parte del resto de la Tierra, lejos de las ciudades de *ems*, se deja a los humanos, ya que los *ems* no están realmente interesados en ella.

Hablando de humanos, querían oír hablar acerca de eso. Los humanos deben retirarse, a la vez y para bien. Simplemente no pueden competir. Ahora, los humanos empiezan a poseer todo el capital de este mundo. La economía crece muy rápido; su riqueza crece muy rápido. Los humanos se hacen ricos colectivamente. Como saben, la mayoría de los humanos hoy día no poseen realmente mucho, aparte de su capacidad para trabajar, por lo que entre ahora y luego, necesitan adquirir suficientes activos, seguros o acuerdos de participación, o pueden morir de hambre. Recomendando encarecidamente evitar este resultado.

Ahora, podrían preguntarse, ¿por qué los *ems* permitirían a los humanos existir? ¿Por qué no matarlos y tomar sus cosas? Pero noten que tenemos muchos jubilados improductivos, que nos rodean hoy día y no los matamos y tomamos sus cosas.

En parte, eso se debe a que rompería las instituciones que compartimos con ellos. Otros grupos se preguntarían quién es el próximo. De manera muy plausible, los *ems* sí pueden permitir a los humanos retirarse en paz durante la era de *ems*. Uds. deberían preocuparse más por la edad de un *em*, ya que solo dura un año o dos y no se sabe lo que sucede a continuación.

Los **ems** son muy parecidos a los humanos, pero no son como el humano típico. El típico **em** es una copia de los pocos cientos de humanos más productivos. De hecho, son como una élite en comparación con el típico humano, como el típico multimillonario, Premio Nobel, medallista olímpico de oro, jefe de Estado. Los **ems** observan a los humanos, quizás con nostalgia y gratitud, pero no tanto con respeto, lo cual es, si piensan en ello, cómo Uds. piensan sobre sus antepasados. Sabemos mucho sobre cómo los humanos difieren en términos de productividad. Solo podemos usarlos para predecir características de **ems**. Por ejemplo, tienden a ser inteligentes, conscientes, trabajadores, casados, religiosos, de mediana edad. Son las características de los **ems**.

El mundo de **ems** también tiene una gran variedad. No solo la mayoría del tipo de variedades humanas, incluyendo la variedad de industria y profesión, también tienen muchos tipos nuevos de variedad, y uno de los más importantes es la velocidad de la mente. Es posible para los **ems** ir desde la velocidad humana, hasta un millón de veces más rápido que la velocidad humana y hasta mil millones de veces más lento que la velocidad humana. Los **ems** más rápidos tienden a tener marcadores de alto estatus. reflejan más riqueza, ganan discusiones, se sientan en lugares privilegiados. Los más lentos son sobre todo jubilados y son como los fantasmas de nuestra literatura. Si recuerdan, los fantasmas nos rodean a todos; pueden interactuar con ellos si pagan el precio, pero no saben mucho, no pueden influir mucho y están obsesionados con el pasado, así que ¿cuál es el punto?

Los **ems** también tienen más variedad en la estructura de sus vidas. Esta es su vida: empiezan y terminan, muy simple. Esta es la vida de un **em** que hace algunas copias a corto plazo cada día, para hacer tareas a corto plazo y luego terminar. Hablaremos más sobre esas versiones a corto plazo en un momento; pero son mucho más eficientes porque no necesitan descansar para el día siguiente. Este **em** es más oportunista. Hacen más copias de sí mismos cuando hay más demanda de eso. No saben de qué manera va el futuro.

Este es un **em** de diseñador que concibe un sistema grande y luego se divide recursivamente en copias que elaboran eso, para que los **ems** puedan implementar diseños más grandes y más coherentes. Este es un plomero de emulación, que recuerda que todos los días, por los últimos 20 años, solo han trabajado dos horas diarias; una vida de ocio. Pero lo que realmente sucedió es que cada día tenían mil copias, y cada una hizo un trabajo de dos horas de plomería y solo una continuó al día siguiente. Objetivamente están trabajando bien más del 99 % del tiempo. Subjetivamente, recuerdan una vida de ocio.

Esto de nuevo, son Uds. Empiezan y terminan. Esto podría ser Uds. si al inicio de la fiesta, tomaran una droga que provocara que no recordaran esa fiesta, absolutamente nada después de ese día. Me dicen que algunas personas hacen esto. Mi pregunta es,

¿al final de la fiesta se dirían a sí mismos: 'Estoy a punto de morir, esto es terrible. Esa persona mañana no seré yo, porque no recordará lo que hago?'. O podrían decir: "Continuaré con el mañana, solo que no recordaré lo que hice". Este es un **em** que divide una copia a corto plazo, para hacer una tarea a corto plazo y luego terminar. Tiene las mismas dos posibles actitudes. Puede decir: "Soy una nueva criatura a corto plazo de corta vida. Odio esto". O: "Soy una parte de una criatura más grande que no recordará esta parte". Predigo que tendrán esa segunda actitud, no porque sea filosóficamente correcto, sino porque les ayuda a avanzar.

Hoy, si el presidente dice que debemos invadir Irak, y Uds. dicen: ¿por qué? y les dicen: secreto de Estado; no están seguros si pueden confiar en ellos, pero los **ems**, copias de Uds. y del presidente pueden ir dentro de una caja fuerte, explicar todas sus razones secretas y luego un bit sale de su copia de sí mismos, diciéndoles si están convencidos. Así que ahora, pueden saber que hay una buena razón.

Sé que están ansiosos por evaluar este mundo. Están ansiosos por decidir si les gusta o lo odian. Pero piensen, sus antepasados de hace miles de años habrían amado u odiado su mundo, con base a las primeras cosas que escucharon sobre él, porque su mundo es realmente extraño. Así que antes de juzgar un extraño mundo futuro, realmente deberían aprender mucho sobre él, quizá leer un libro al respecto y luego, si no les gusta: trabajar para cambiarlo.

Gracias.